How Aberrant *N*-Glycosylation Can Alter Protein Functionality and Ligand Binding: an Atomistic View

Matteo Castelli[1], Pengrong Yan[2], Anna Rodina[2], Chander S. Digwal[2], Palak Panchal[2], Gabriela Chiosis[2,3,*], Elisabetta Moroni[4,*], Giorgio Colombo[1,5*]

1. Department of Chemistry, University of Pavia, via Taramelli 12, 27100 Pavia, Italy.
2. Chemical Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
3. Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA.
4. SCITEC-CNR, via Mario Bianco 9, 20131 Milano, Italy.
5. Lead contact

*Correspondence: g.colombo@unipv.it (G.Colombo.); elisabetta.moroni@scitec.cnr.it (E.M.); chiosisg@mskcc.org (G.Chiosis.)

Supplementary Information

Supplementary Figure 1

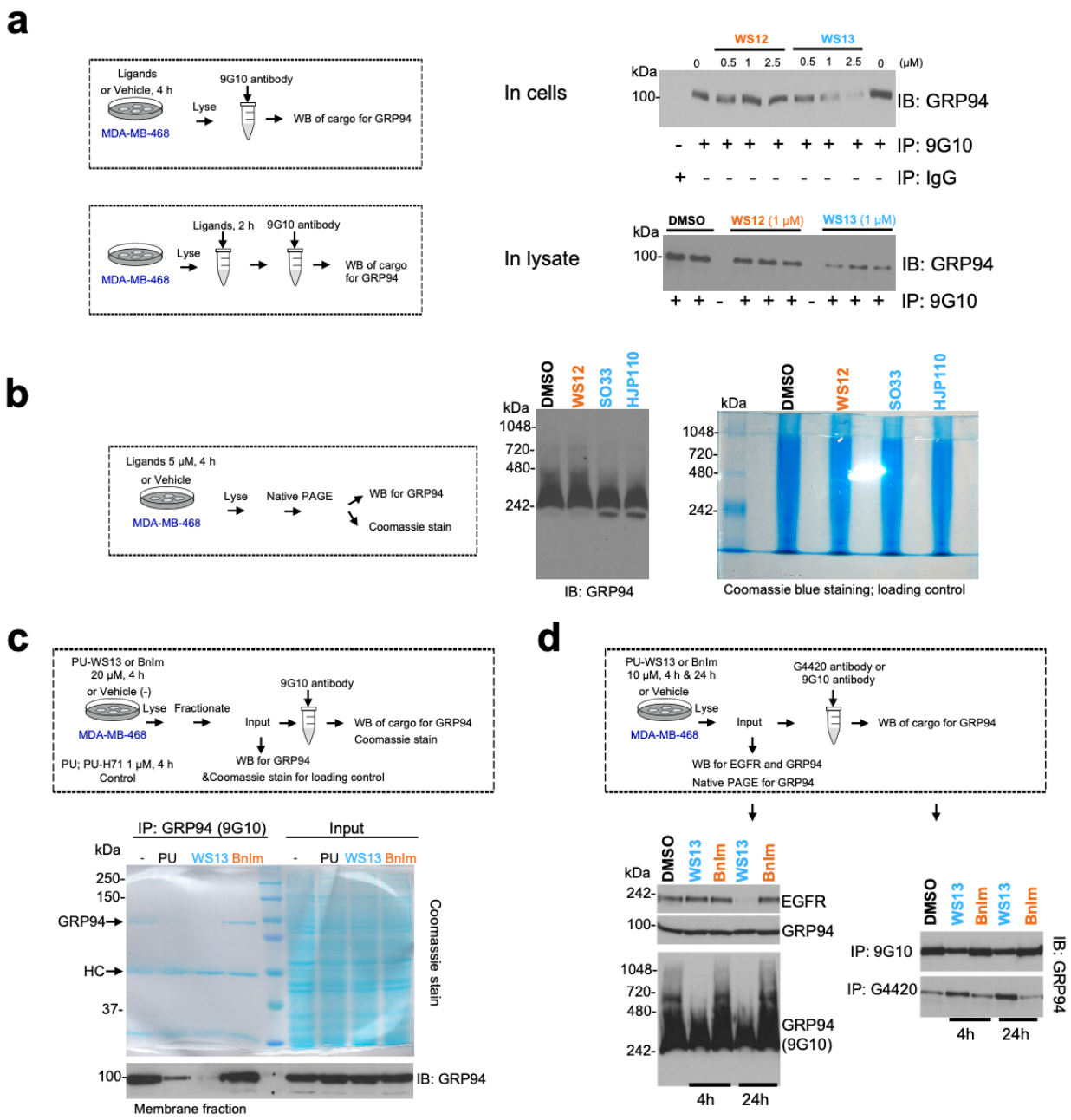


Figure S1. Testing paradigm for small molecule activity via ^{Glyc62}GRP94, Related to Figure 1.
(a) Immuno-capture of GRP94 was performed as shown in the schematic. MDA-MB-468 cells or cell homogenates were treated with ligands or vehicle (DMSO) as indicated. Top panel, for 4 h at 0.5, 1 and 2.5 μM and lower panel, for 2 h at 1 μM, prior to immunocapture and WB of the cargo. Experiments were repeated thrice with similar results. IgG, isogenic control.
(b) Native-PAGE followed by immunoblotting for GRP94 (left) or Coomassie staining (right). MDA-MB-468 cancer cells were treated with vehicle (DMSO) or indicated ligands (at 5 μM for 4 h) prior to lysing and analysis. Gels are representative of three independent results.
(c) As in (a) for MDA-MB-468 cells treated with ligands (at 20 μM for 4 h) or vehicle (DMSO) prior to immunocapture and WB of the cargo. The membrane fraction was used for immunocapture. The WB and Coomassie stained gels of the input are shown as control for equal protein amount, as used for immunocapture. Gels are representative of three independent results.
(d) Functional and biochemical analysis of PU-WS13 and BnIm in a battery of assays that test productive ^{Glyc62}GRP94 engagement in cells. MDA-MB-468 cells were treated for 4h or 24 h with inhibitors (at 10 μM) prior to analysis, as indicated in the schematic. ^{Glyc62}GRP94 inhibits EGFR internalization and degradation, and ^{Glyc62}GRP94 inhibition suppresses this effect (see EGFR steady-state level analysis by WB). GRP94 conformational modulation by ligands was detected by immunocapture and Native PAGE, as indicated.

Supplementary Figure 2

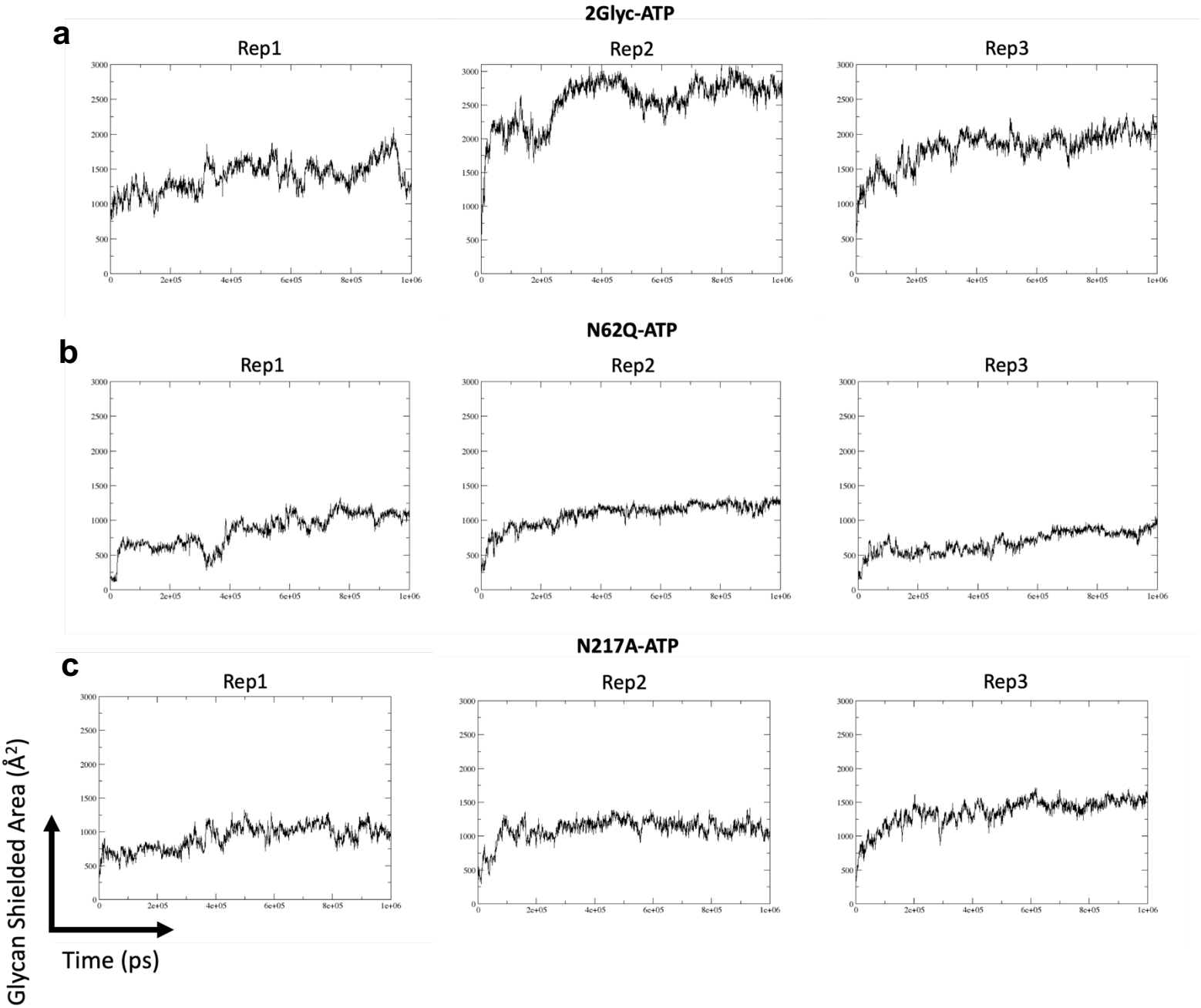


Figure S2. Glycan shielded area for GRP94 variants over the entire single replica, Related to Figure 3.
(a) Analysis for the 2Glyc-ATP system.
(b) Analysis for the N62Q-ATP system.
(c) Analysis for the N217A-ATP system.

Supplementary Figure 3

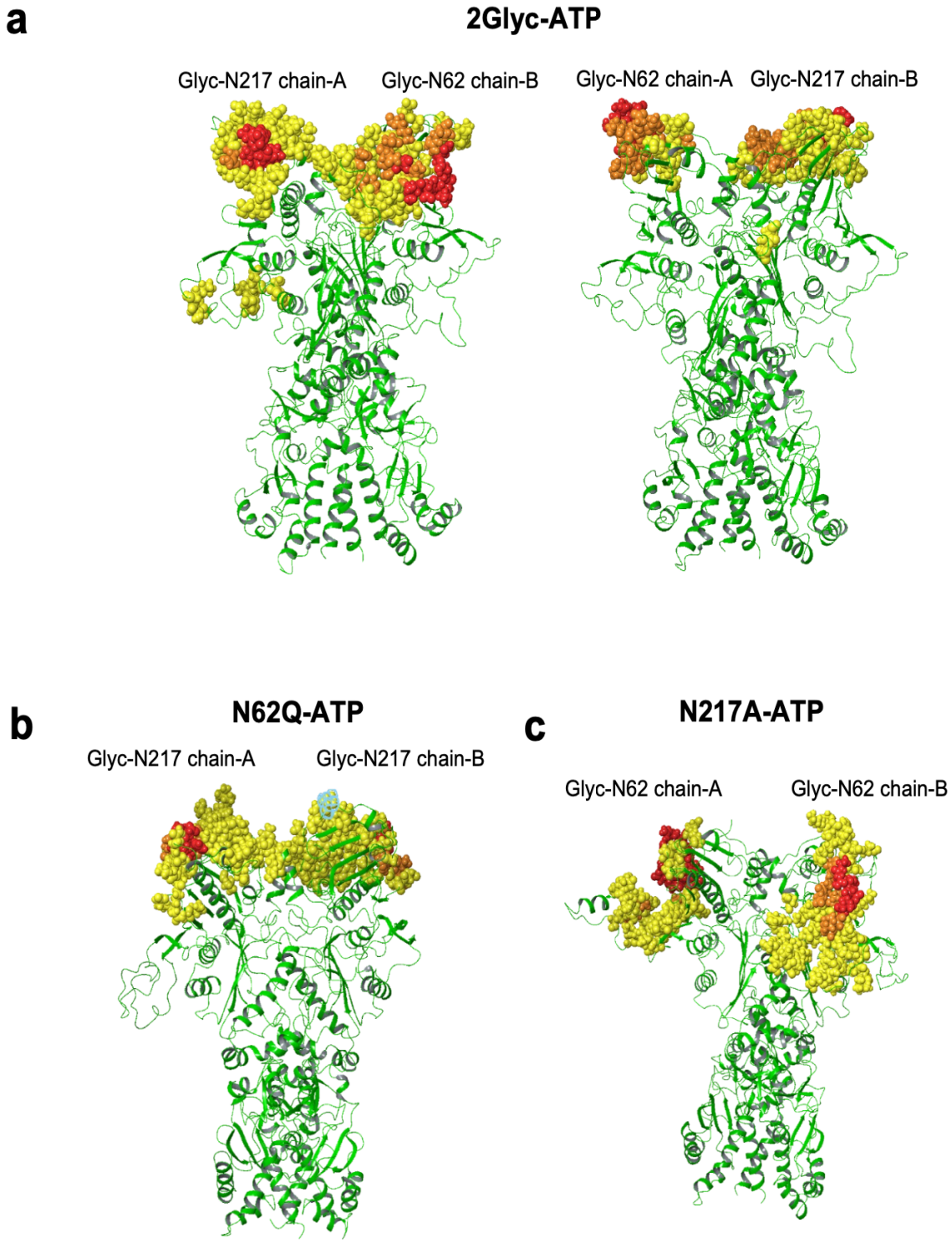


Figure S3. Binding of ATP to the GRP94 variants, Related to Figure 3.
(a) Contact between sugar chain and protein for 2Glyc-ATP.
(b) Contact between sugar chain and protein for N62Q-ATP.
(c) Contact between sugar chain and protein for N217A-ATP.
In yellow we represent residues of the proteins that make contact with the sugar 20% to 50% of the time, orange 50% to 70%, red 70% to 100%.
The data from the simulations are provided electronically as a .zip file (Data_Glyc_Contacts.zip).

Supplementary Figure 4

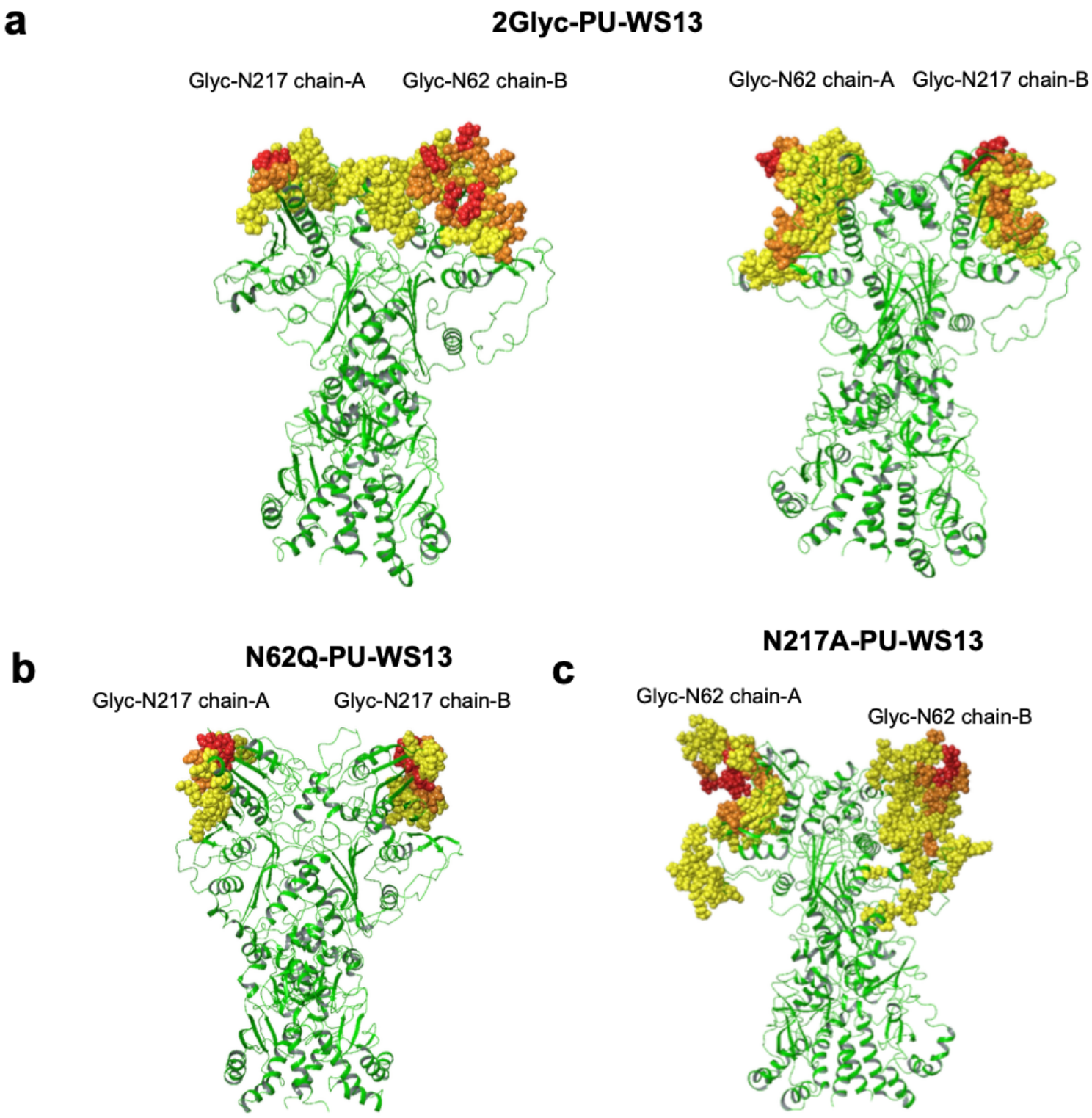


Figure S4. Binding of PU-WS13 to the GRP94 variants, Related to Figure 3.
(a) Contact between sugar chain and protein for 2Glyc-PU-WS13.
(b) Contact between sugar chain and protein for N62Q-PU-WS13.
(c) Contact between sugar chain and protein for N217A-PU-WS13.
In yellow we represent residues of the proteins that make contact with the sugar 20% to 50% of the time, orange 50% to 70%, red 70% to 100%.
The data from the simulations are provided electronically as a .zip file (Data_Glyc_Contacts.zip).

Supplementary Figure 5

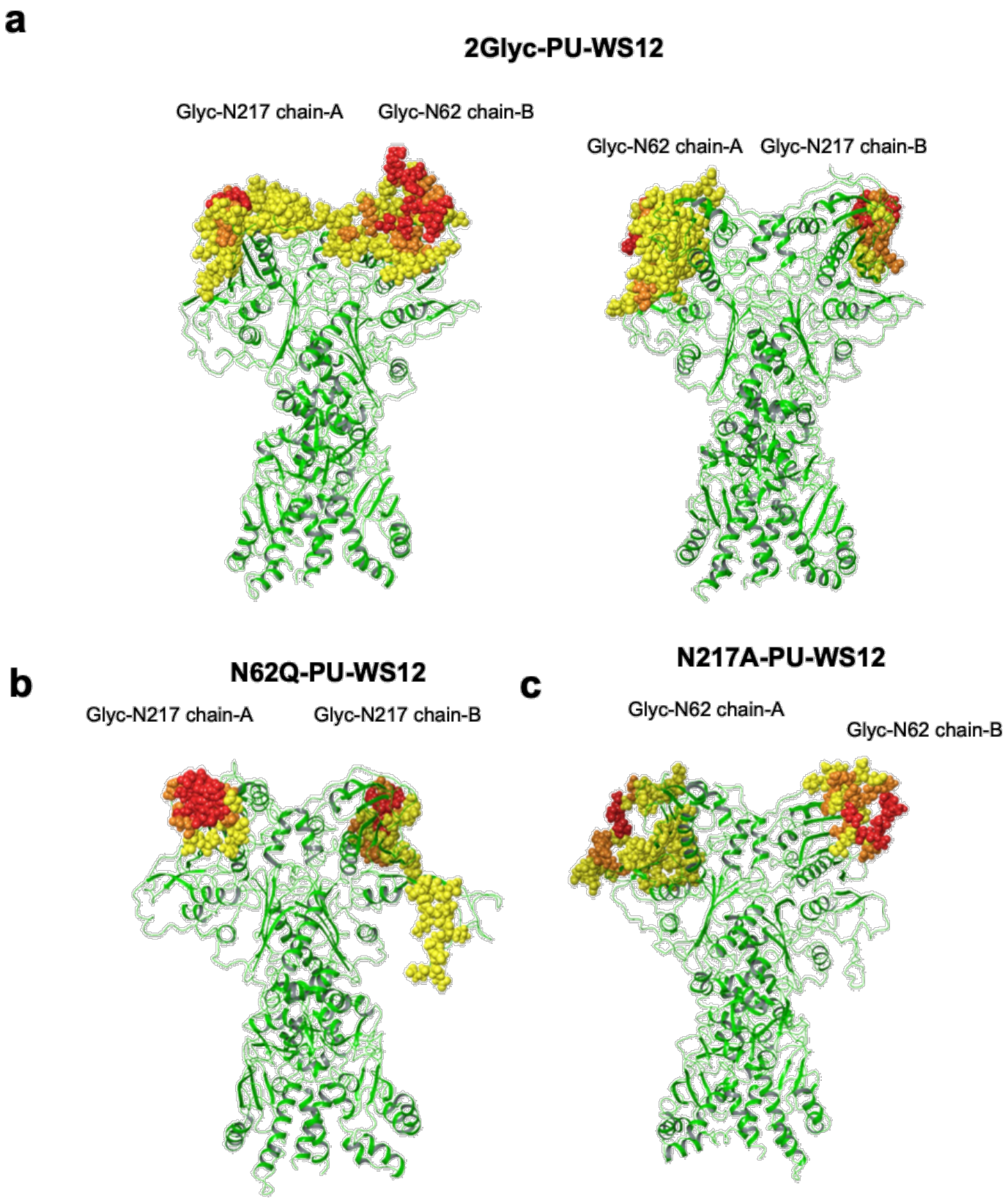


Figure S5. Binding of PU-WS12 to the GRP94 variants, Related to Figure 3.
(a) Contact between sugar chain and protein for 2Glyc-PU-WS12.
(b) Contact between sugar chain and protein for N62Q-PU-WS12.
(c) Contact between sugar chain and protein for N217A-PU-WS12.
In yellow we represent residues of the proteins that make contact with the sugar 20% to 50% of the time, orange 50% to 70%, red 70% to 100%.
The data from the simulations are provided electronically as a .zip file (Data_Glyc_Contacts.zip).

Supplementary Figure 6

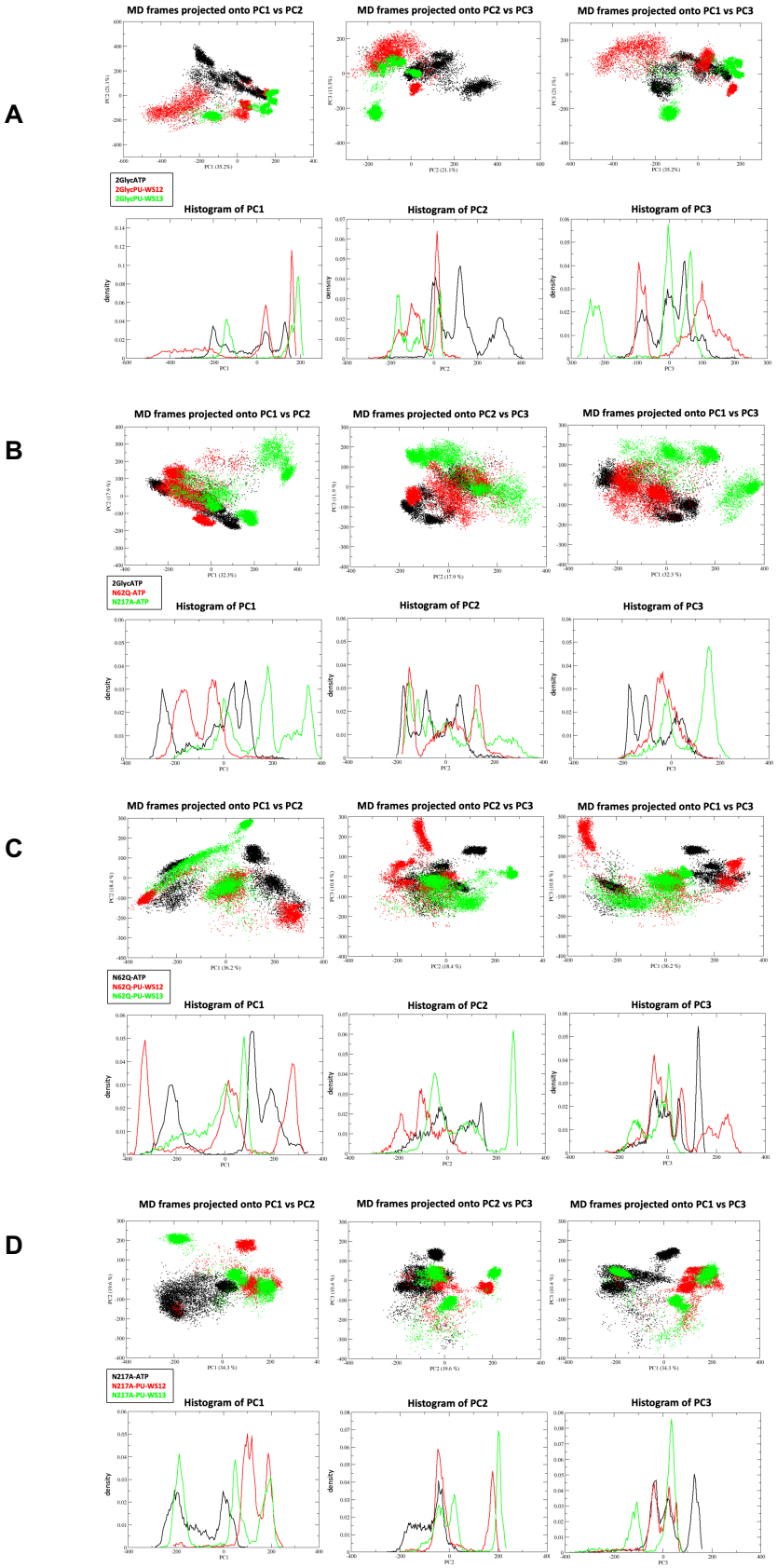


Figure S6. PCA long charged loop, Related to Figure 3. In each subpanel, the top row reports the projection of the dynamics of the charged linker along the 3 principal eigenvectors, while the bottom row reports the histograms of the values of the three principal components. **A)** Simulation 2GlycATP with different ligands. Black is ATP, red PU-WS12, green PU-WS13; **B)** Same as **A,** but with system N62Q; **C)** Same as **A**, but with system N217A; **D)** Comparison of the ATP-states in three different glycosylation states: black is 2GlycATP; red is N62Q; green is N217A.

Supplementary Figure 7

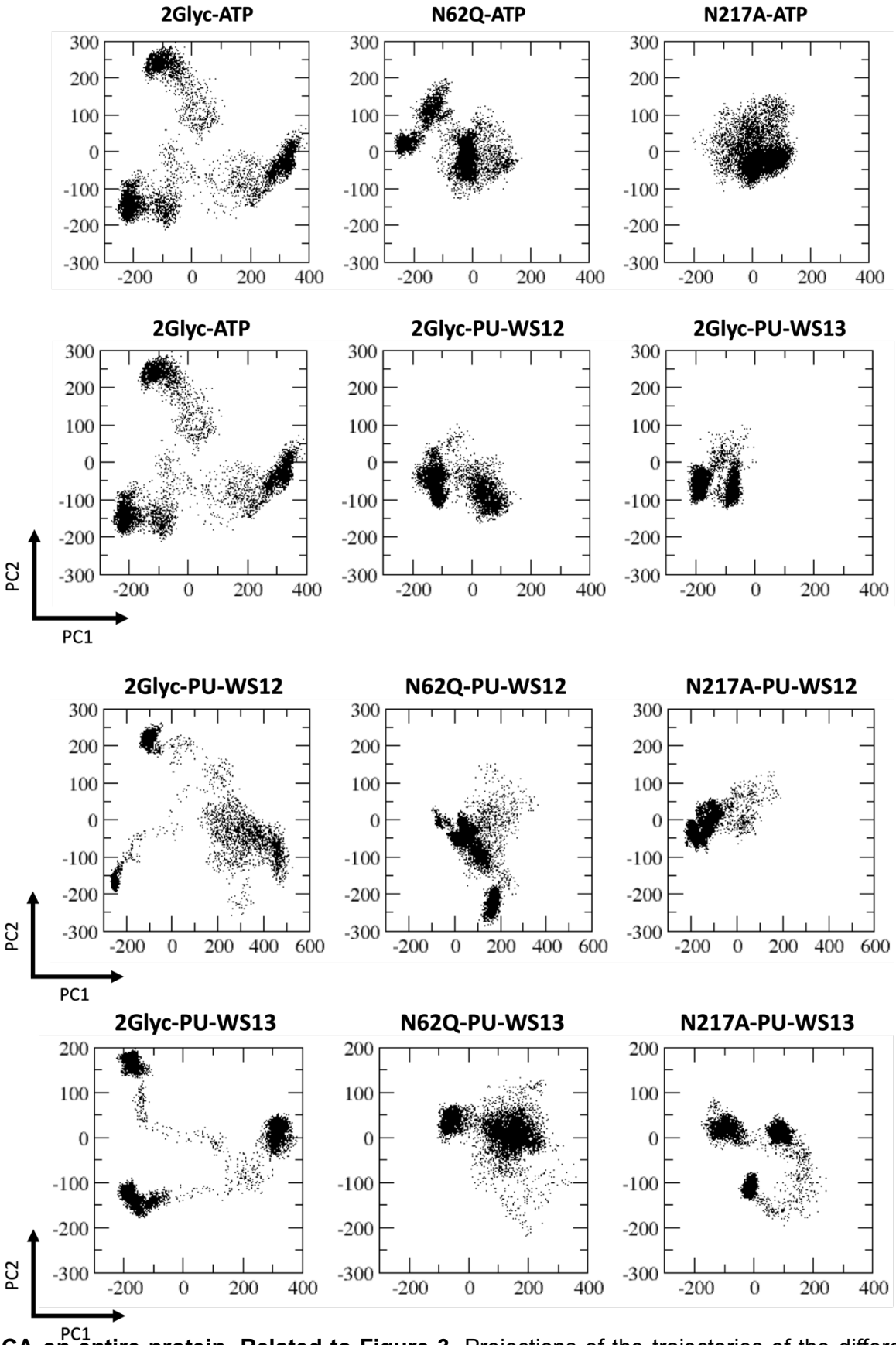


Figure S7. PCA on entire protein, Related to Figure 3. Projections of the trajectories of the different systems simulated on the space defined by the two principal components of the 2GlycATP system. The title above each graph describes the system under exam.

Supplementary Figure 8

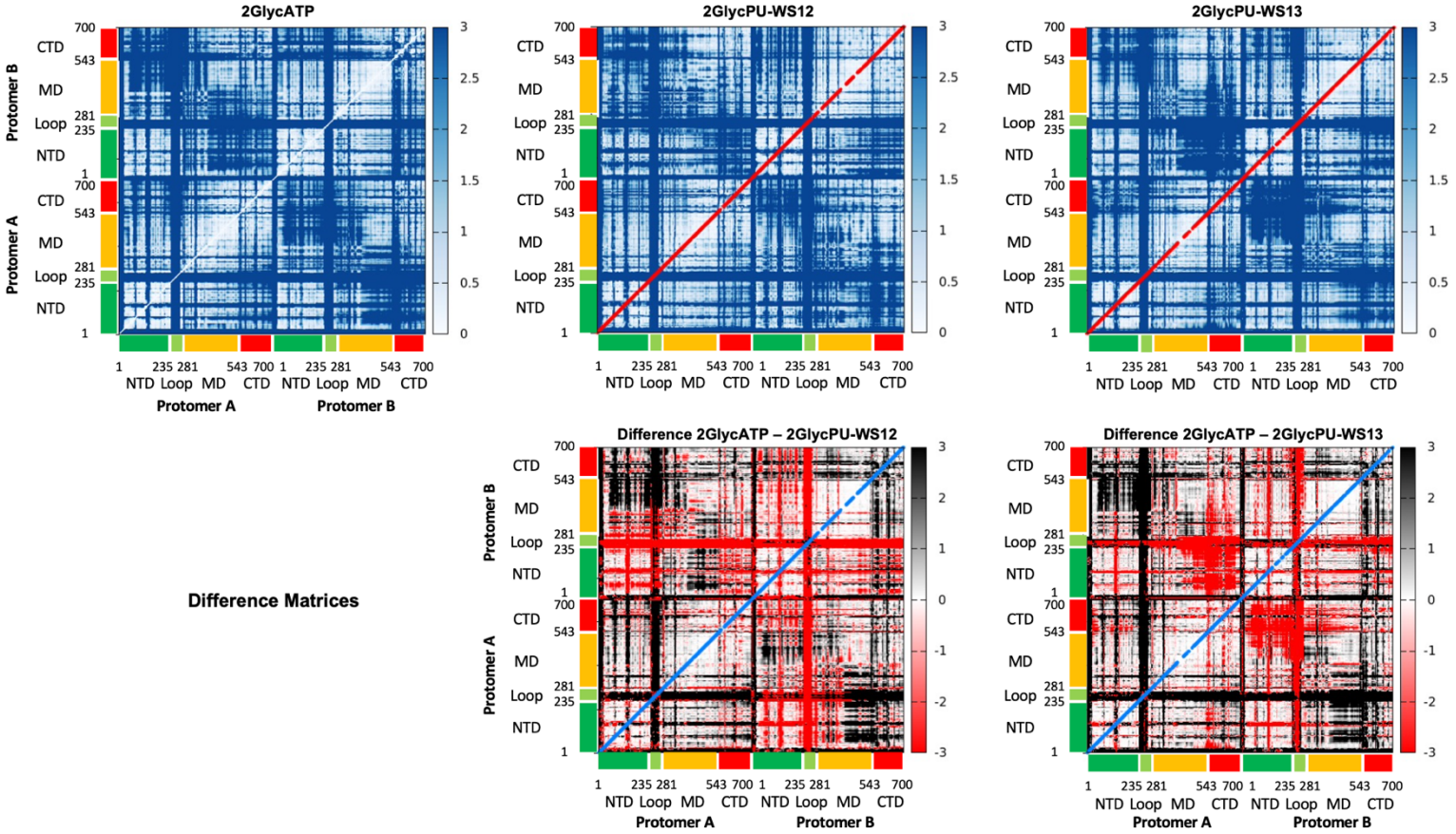


Figure S8. Distance Fluctuation (DF) matrix for the 2Glyc state, bound to ATP, PU-WS12 or PU-WS13, Related to Figures 4 and 5.

The reference DF matrix for the fully glycosylated ATP-state of GRP94 and the difference-matrices obtained by subtracting the DF values of another state (different ligand, mutant, or mutant+ligand) from this reference, are presented. This type of representation is intended to highlight the regions where the different chemical perturbations determine the largest impact on the internal dynamics of the protein: the color code indicates lower coordination (higher dynamic flexibility, less ordered motions, easier disruption of local interaction networks) with respect to the reference state in red, and higher coordination (increased communication between different areas) in black. Matrix representation of the residue-pair DFs was calculated from the various trajectories. In each subpanel, the top row represents the original matrix, while the bottom row represents the difference matrix obtained subtracting the values for a certain mutant/ligand state from the reference matrix corresponding to the 2GlycATP state. If residue-pair fluctuations are larger (lower coordination, larger flexibility) in the specific mutant/ ligand state than in the reference state, the corresponding value in the difference is depicted in red; if residue- pair fluctuations are smaller (higher coordination, lower flexibility) in the specific mutant/ligand state than in the reference state, the corresponding value in the difference is depicted in black.

Supplementary Figure 9

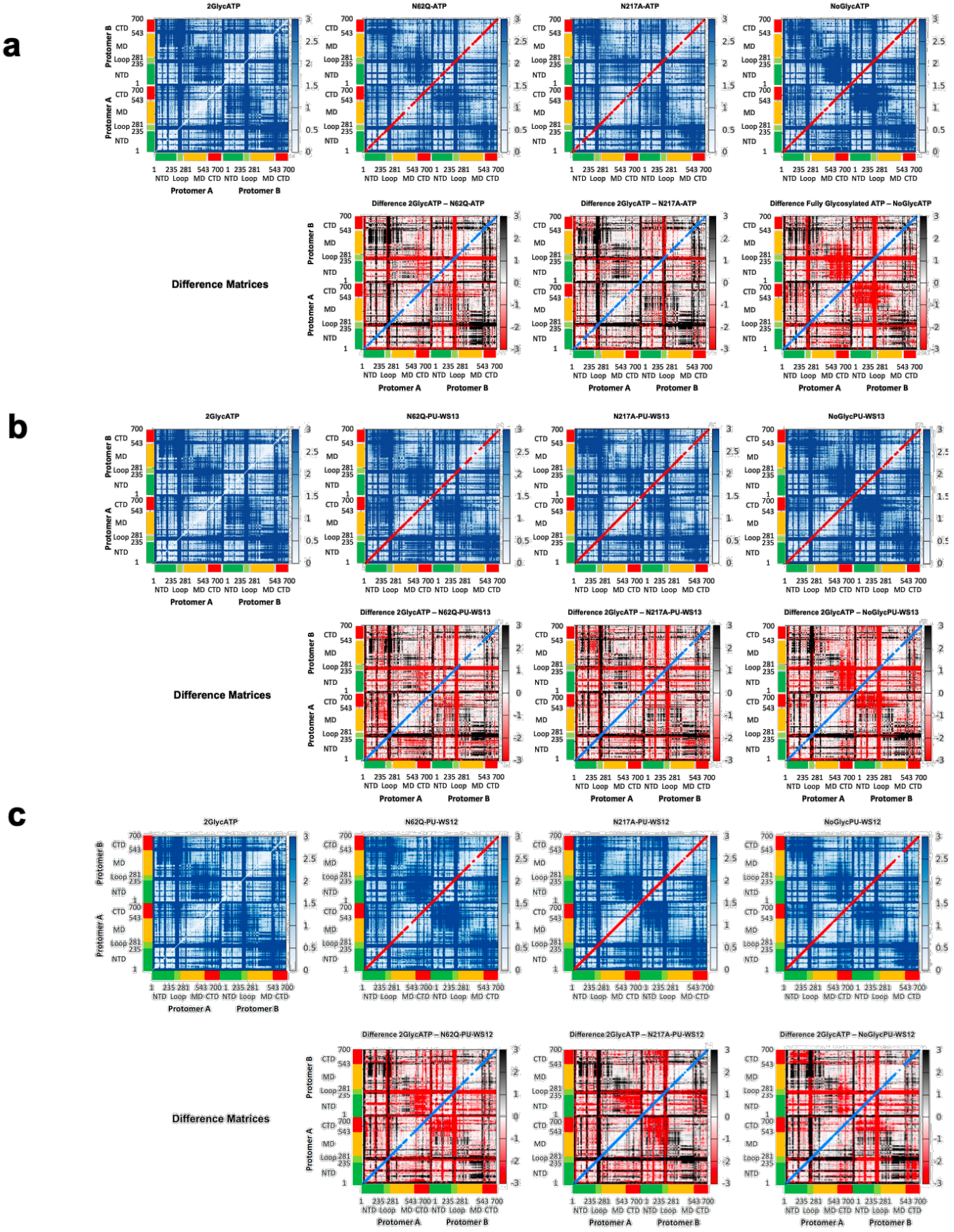


Figure S9. Distance Fluctuation matrix for different mutants, including the unglycosylated Wild Type, in complex with various ligands, Related to Figures 4 and 5.

(a) DF matrix for GRP94 variants in complex with ATP.
(b) DF matrix for GRP94 variants in complex with PU-WS13.
(c) DF matrix for GRP94 variants in complex with PU-WS12.

Supplementary Figure 10

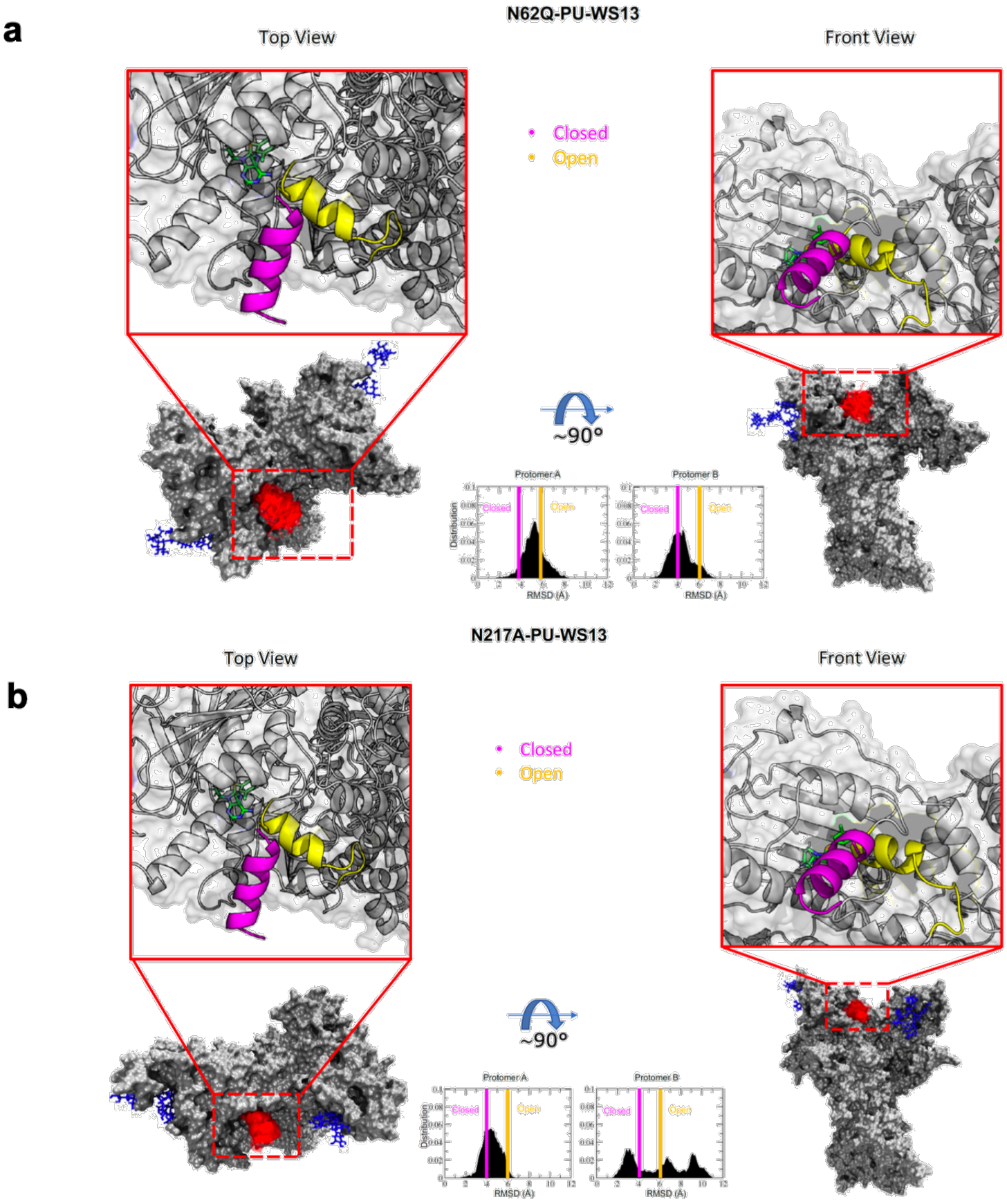


Figure S10. Conformational dynamics of the ATP-lid covering the active site of GRP94 variants and their interactions with sugar chains in the open or closed forms, Related to Figures 7, 8 and 9.
(a) Analyses for the N62Q-PU-WS13 system.
(b) Analyses for the N217A-PU-WS13 system.
Two orientations are depicted. ATP-lid (at several times, 300 frames) and glycans are represented with red and blue lines, respectively. The insets show the distributions of RMSD values from the conformation presented in crystal structure PDB 5ULS.pdb. The blue and yellow vertical lines show the RMSD threshold for the closed and open states, respectively.

Supplementary Figure 11

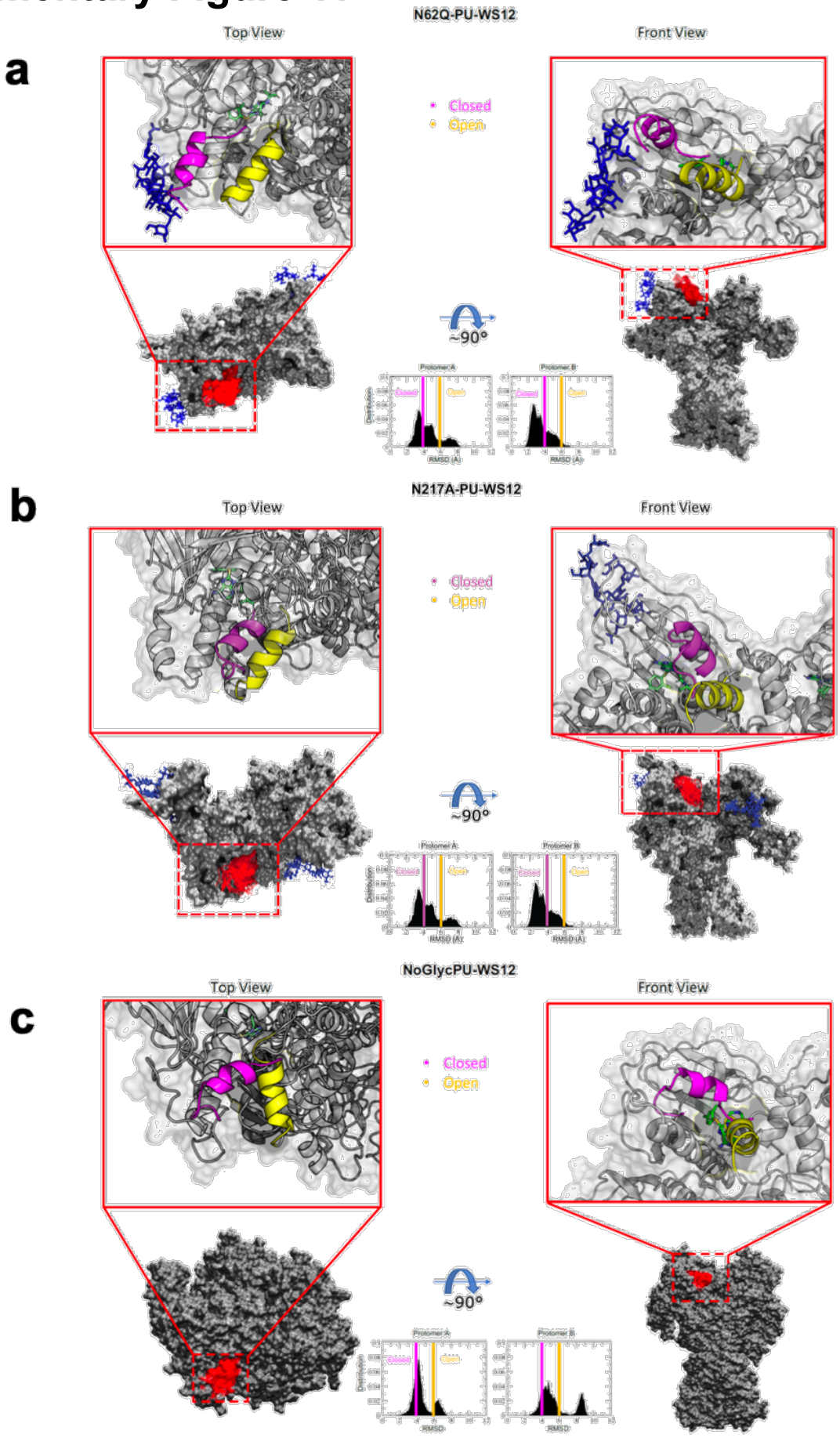


Figure S11. Conformational dynamics of the ATP-lid covering the active site of GRP94 variants and their interactions with sugar chains in the open or closed forms, Related to Figures 7, 8 and 9.
(a) Analyses for the N62Q-PU-WS12 system.
(b) Analyses for the N217A-PU-WS12 system.
(c) Analyses for the NoGlyc-PU-WS12 system.
Two orientations are depicted. ATP-lid (at several times, 300 frames) and glycans are represented with red and blue lines, respectively. The insets show the distributions of RMSD values from the conformation presented in crystal structure PDB 5ULS.pdb. The blue and yellow vertical lines show the RMSD threshold for the closed and open states, respectively.

Supplementary Figure 12

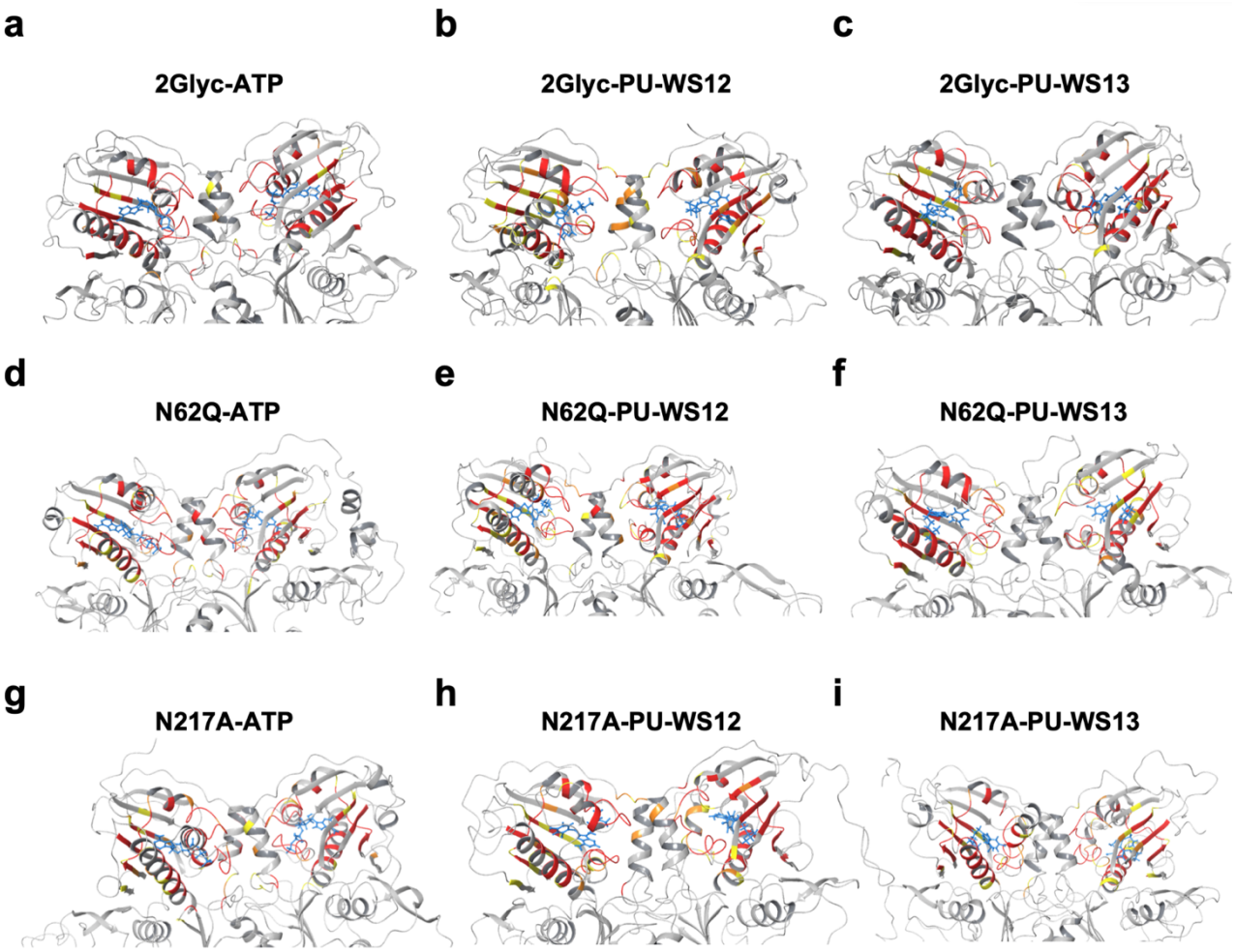


Figure S12. **Contact between ligand and surrounding residues of the protein for GRP94 variants in complex with ligands, Related to Figure 9.**
(a) The 2Glyc-ATP system.
(b) The 2Glyc-PU-WS12 system.
(c) The 2Glyc-PU-WS13 system.
(d) The N62Q-ATP system.
(e) The N62Q-PU-WS12 system.
(f) The N62Q-PU-WS13 system.
(g) The N217A-ATP system.
(h) The N217A-PU-WS12 system.
(i) TheN217A-PU-WS13system.
In yellow we represent residues of the proteins that make contact with the ligand 20% to 50% of the time, orange 50% to 70%, red 70% to 100%.
The data from the simulations are provided electronically as a .zip file (Data_Ligands_Contacts.zip).

Supplementary Figure 13

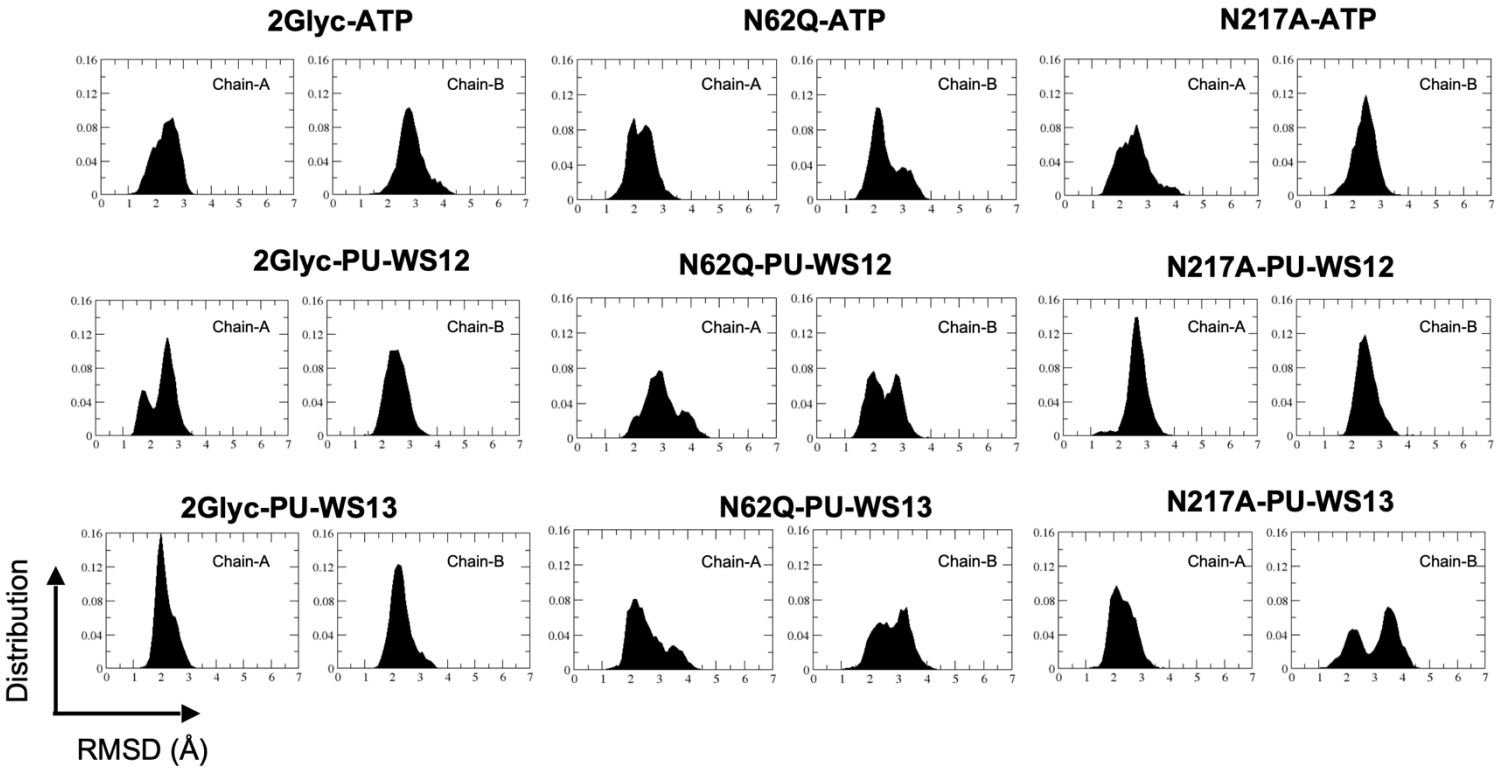


Figure S13. RMSD distribution over the entire metatrajectory for indicated GRP94 variant-ligand systems, Related to Figure 9.

Calculation was carried out independently for each protomer and focused on the residues making up the active

Supplementary Figure 14

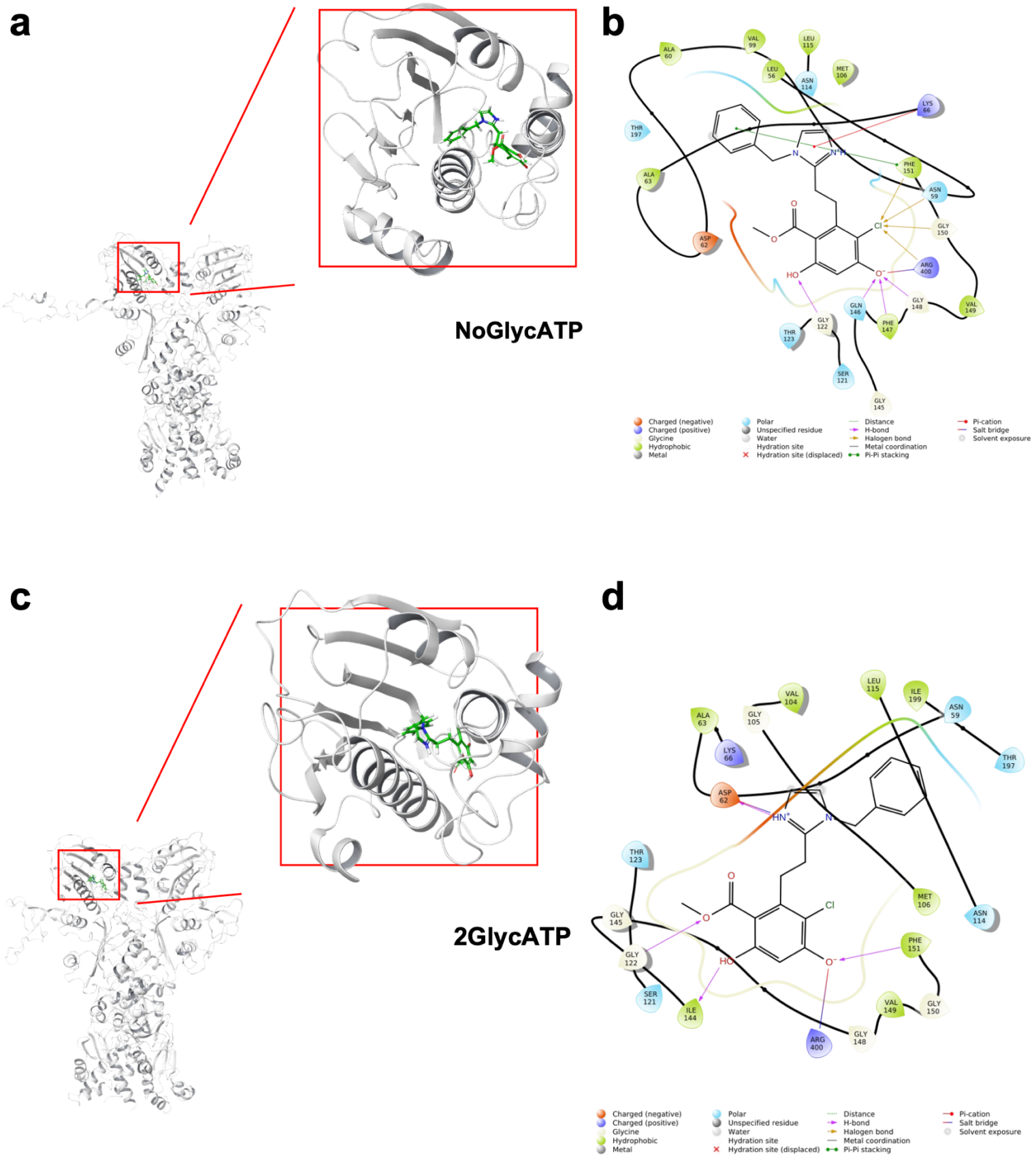


Figure S14. Bnlm docking analysis in GRP94 variants, Related to Figure 2.
(a) Representative structure for the best scoring docking pose of Bnlm into the active site of GRP94 NoGlycATP. Zoom on NTD.
(b) 2D-interaction di agram based on the best-scoring docking pose of Bnlm as in (a). Figure generated with Schrodinger's Maestro suite.
(c) Representative structure for the best scoring docking pose of Bnlm into the active site of GRP94 2GlhycATP. Zoom on NTD.
(d) 2D-interaction diagram based on the best-scoring docking pose of Bnlm as in (c).
See Supplementary data provided electronically (Data_Docking_Bnlm.xlsx)

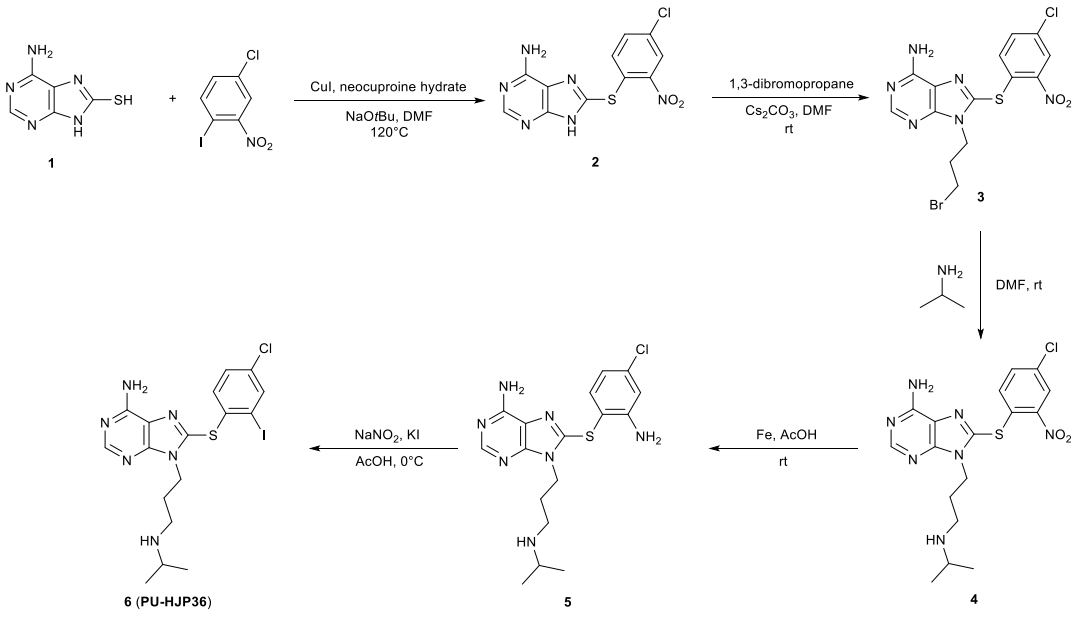
Data S1. Analysis of accessible area for more than 70% of the simulation time, Related to Figure 3 insets.

2Glyc-ATP:
accessible > 70 %: 1, 2, 4, 5, 7, 9-15, 17, 18, 20, 21, 23, 26, 31, 32, 34, 35, 37-75, 77-119, 121-127, 129-159, 161-165, 174-182, 184, 186-190, 196-723, 750-754, 756, 759, 762, 763, 766-787, 789-791, 793-795, 802-825, 827, 833, 837-841, 855-873, 875, 876, 878, 879, 884-887, 896, 904, 906, 907, 909-916, 918,925, 927-1446

N62Q-ATP:
accessible > 70 %: 1-75, 77, 78, 83-106, 114-165, 174-181, 187-189, 197-711, 713, 714, 717-733, 740-777, 779, 780, 782-784, 792-817, 822, 823, 826-830, 832-835, 838, 839, 842-876, 883, 885-905, 907-1424

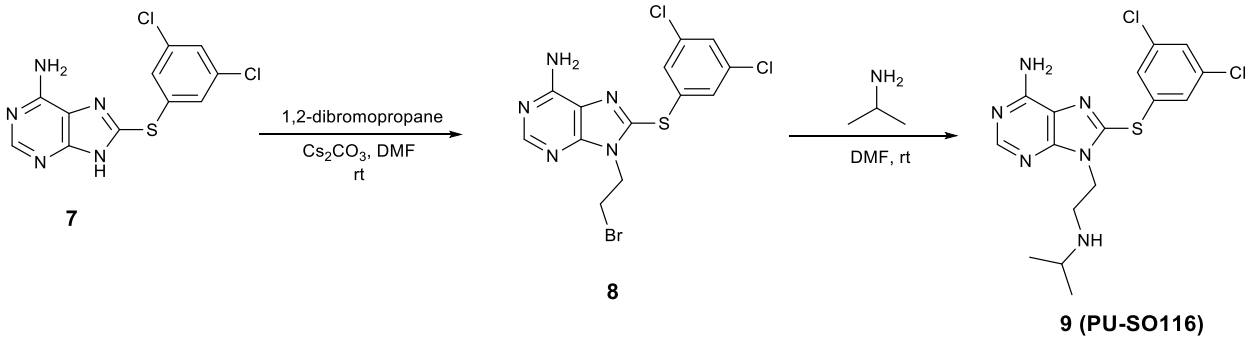
N217A-ATP:
accessible > 70 %: 1, 2, 4-7, 16, 18, 26-88, 90, 92, 95, 97-188, 190-233, 235, 237-245, 255-279, 285-717, 734-759, 761

Supplementary Scheme 1



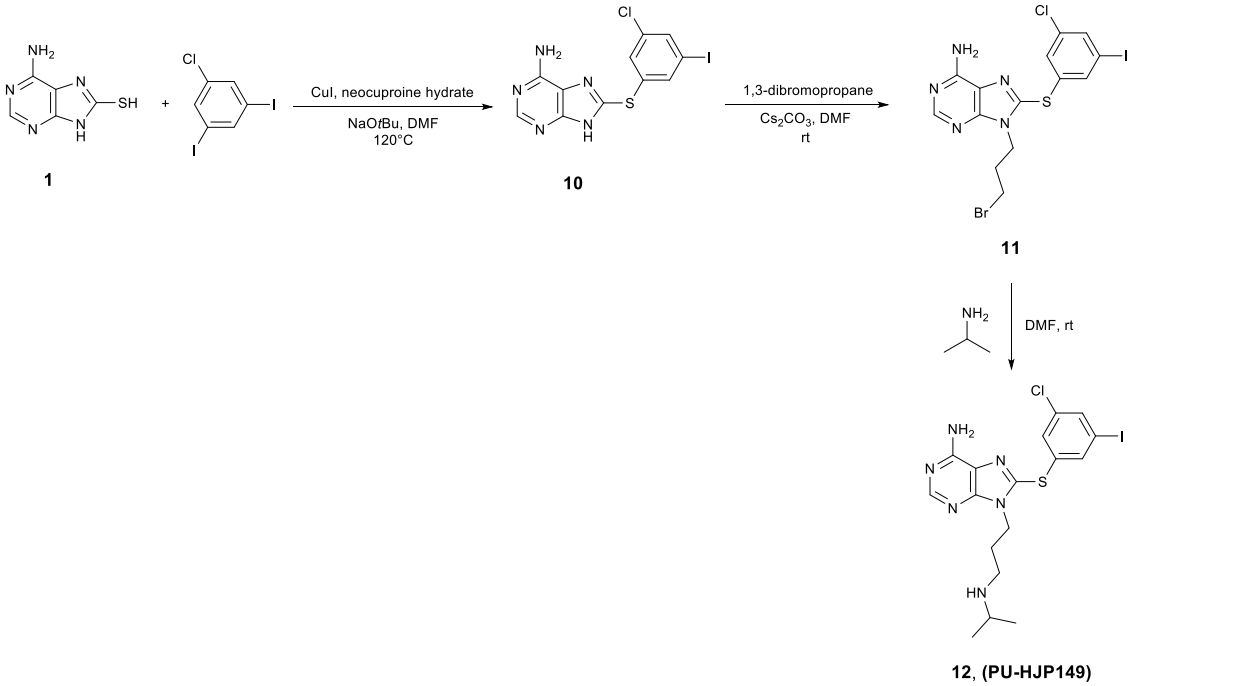
Scheme S1. Synthetic scheme of PU-HJP36, Related to STAR Methods.

Supplementary Scheme 2



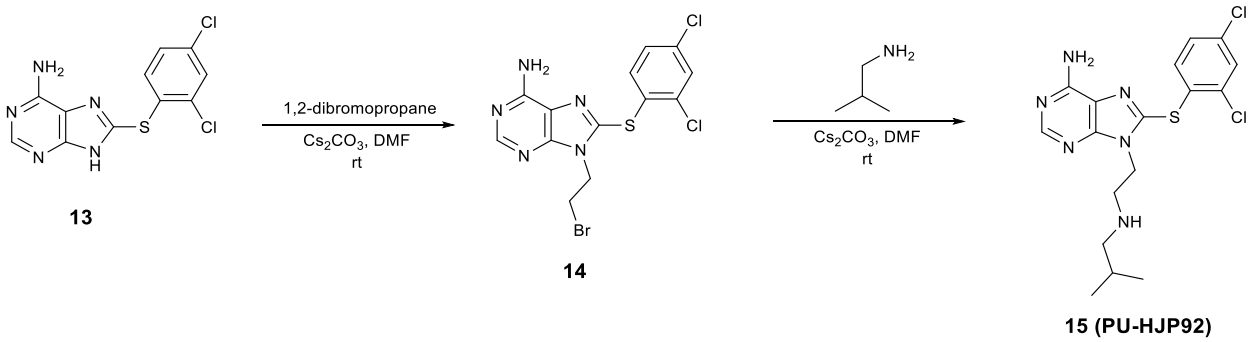
Scheme S2. Synthetic scheme of PU-SO116, Related to STAR Methods.

Supplementary Scheme 3



Scheme S3. Synthetic scheme of PU-HJP149, Related to STAR Methods.

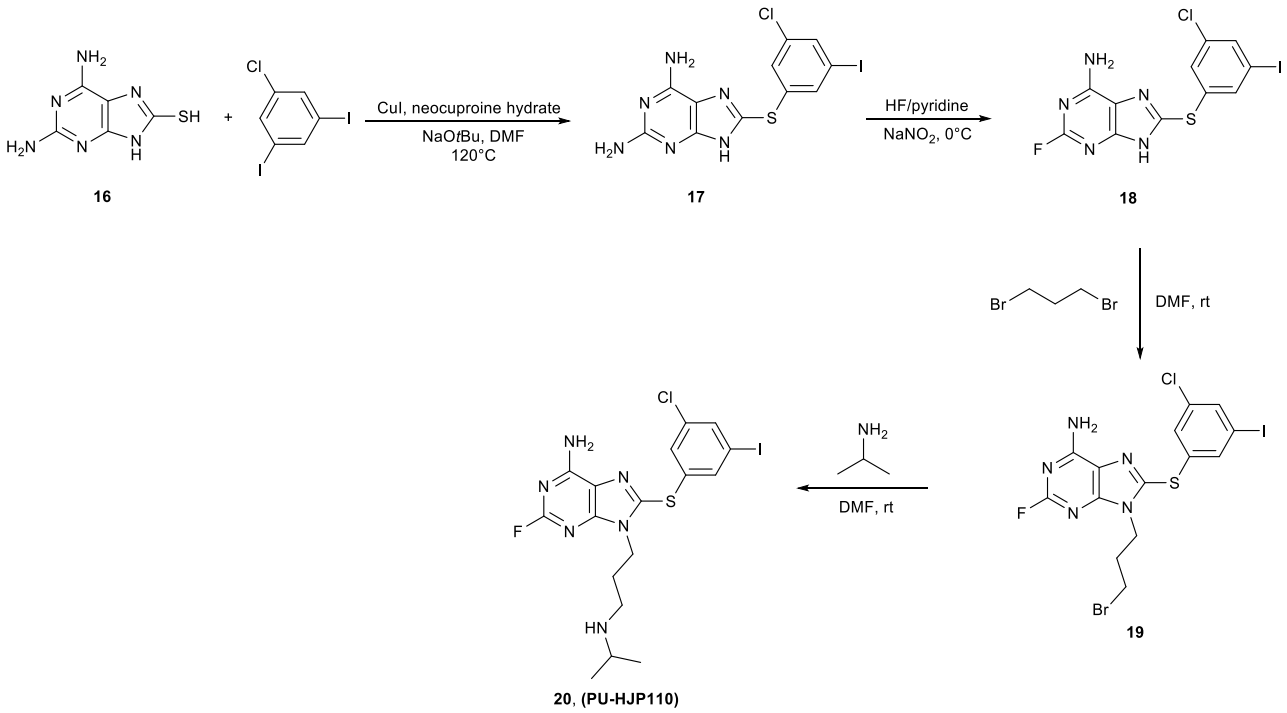
Supplementary Scheme 4



13 14 15 (PU-HJP92)

Scheme S4. Synthetic scheme of PU-HJP92. Related to STAR Methods.

Supplementary Scheme 5



16 17 18 19 20, (PU-HJP110)

Scheme S5. Synthetic scheme of PU-HJP110, Related to STAR Methods.